AD NO. ______________________
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC-9087

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

MOGULS SCORING RECORD NO. 669

SITE LOCATION:
U.S. ARMY YUMA PROVING GROUND

DEMONSTRATOR:
NAEVA GEOPHYSICS INC.
P.O. BOX 7325
CHARLOTTESVILLE, VA 22906

TECHNOLOGY TYPE/PLATFORM:
EM61 MKII/MAN-PORTABLE

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

SEPTEMBER 2005

Prepared for:
U.S. ARMY ENVIRONMENTAL CENTER
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISPOSITION INSTRUCTIONS

Destroy this document when no longer needed. Do not return to the originator.

The use of trade names in this document does not constitute an official endorsement or approval of the use of such commercial hardware or software. This document may not be cited for purposes of advertisement.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
September 2005	Final	7, 13, 14, and 16 December 2004

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE MOGULS SCORING RECORD NO. 669 (NAEVA GEPHYSICS, INC.)	
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Overbay, Larry; Robitaille, George The Standardized UXO Technology Demonstration Site Scoring Committee	8-CO-160-UXO-021
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-AT-SL-E Aberdeen Proving Ground, MD 21005-5059	ATC-9087
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
Commander U.S. Army Environmental Center ATTN: SFIM-AEC-ATT Aberdeen Proving Ground, MD 21005-5401	
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This scoring record documents the efforts of NAEVA Geophysics, Inc., to detect and discriminate inert unexploded ordnance (UXO) utilizing theYPG Standardized UXO Technology Demonstration Site Moguls. Scoring Records have been coordinated by Larry Overbay and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include, the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Center, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS
NAEVA Geophysics, Inc., UXO Standardized Technology Demonstration Site Program, Moguls EM61 MKII/hand held

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
Unclassified	Unclassified	Unclassified	UL		19b. TELEPHONE NUMBER *(Include area code)*

ACKNOWLEDGEMENTS

Authors:

Larry Overbay Jr.
Matthew Boutin
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground (APG)

Robert Archiable
EC 111, Limited Liability Company (LLC)
U.S. Army Yuma Proving Ground (YPG)

Christina McClung
Aberdeen Test and Support Services (ATSS)
Tri-s, Inc., Aberdeen Data Services Team (ADST)
U.S. Army Aberdeen Proving Ground (APG)

Contributor:

George Robitaille
U.S. Army Environmental Center (AEC)
U.S. Army Aberdeen Proving Ground (APG)

TABLE OF CONTENTS

SECTION 4. TECHNICAL PERFORMANCE RESULTS

SECTION 5. ON-SITE LABOR COSTS

SECTION 6. COMPARISON OF RESULTS TO OPEN FIELD DEMONSTRATION

SECTION 7. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of unexploded ordnance (UXO) require testing so that their performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in ordnance and clutter. Testing at these sites is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments.

The Standardized UXO Technology Demonstration Site Program is a multi-agency program spearheaded by the U.S. Army Environmental Center (AEC). The U.S. Army Aberdeen Test Center (ATC) and the U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) provide programmatic support. The program is being funded and supported by the Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP) and the Army Environmental Quality Technology Program (EQT).

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios that vary targets, geology, clutter, topography, and vegetation.

b. To determine cost, time, and manpower requirements to operate the technology.

c. To determine demonstrator's ability to analyze survey data in a timely manner and provide prioritized "Target Lists" with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth, geo-referenced data for post-demonstration analysis.

1.2.1 Scoring Methodology

a. The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating

characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}), and those that do not correspond to any known item, termed background alarms.

b. The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the blind grid RESPONSE STAGE, the demonstrator provides the scoring committee with a target response from each and every grid square along with a noise level below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, will include signals both above and below the system noise level.

c. The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such and to reject clutter. For the blind grid DISCRIMINATION STAGE, the demonstrator provides the scoring committee with the output of the algorithms applied in the discrimination-stage processing for each grid square. The values in this list are prioritized based on the demonstrator's determination that a grid square is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking is based on human (subjective) judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance, (i.e. that is expected to retain all detected ordnance and rejects the maximum amount of clutter).

d. The demonstrator is also scored on EFFICIENCY and REJECTION RATIO, which measures the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. EFFICIENCY measures the fraction of detected ordnance retained after discrimination, while the REJECTION RATIO measures the fraction of false alarms rejected. Both measures are defined relative to performance at the demonstrator-supplied level below which all responses are considered noise, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

e. Based on configuration of the ground truth at the standardized sites and the defined scoring methodology, there exists the possibility of having anomalies within overlapping halos and/or multiple anomalies within halos. In these cases, the following scoring logic is implemented:

(1) In situations where multiple anomalies exist within a single R_{halo}, the anomaly with the strongest response or highest ranking will be assigned to that particular ground truth item.

(2) For overlapping R_{halo} situations, ordnance has precedence over clutter. The anomaly with the strongest response or highest ranking that is closest to the center of a particular ground truth item gets assigned to that item. Remaining anomalies are retained until all matching is complete.

(3) Anomalies located within any R_{halo} that do not get associated with a particular ground truth item are thrown out and are not considered in the analysis.

f. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 3.1.1.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response Stage ROC curves:

(1) Probability of Detection (P_d^{res}).

(2) Probability of False Positive (P_{fp}^{res}).

(3) Background Alarm Rate (BAR^{res}) or Probability of Background Alarm (P_{BA}^{res}).

b. Discrimination Stage ROC curves:

(1) Probability of Detection (P_d^{disc}).

(2) Probability of False Positive (P_{fp}^{disc}).

(3) Background Alarm Rate (BAR^{disc}) or Probability of Background Alarm (P_{BA}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False Positive Rejection Rate (R_{fp}).

(3) Background Alarm Rejection Rate (R_{BA}).

d. Other:

(1) Probability of Detection by Size and Depth.

(2) Classification by type (i.e., 20-, 40-, 105-mm, etc.).

(3) Location accuracy.

(4) Equipment setup, calibration time and corresponding man-hour requirements.

(5) Survey time and corresponding man-hour requirements.

(6) Reacquisition/resurvey time and man-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

1.3 STANDARD AND NONSTANDARD INERT ORDNANCE TARGETS

The standard and nonstandard ordnance items emplaced in the test areas are listed in Table 1. Standardized targets are members of a set of specific ordnance items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are inert ordnance items having properties that differ from those in the set of standardized targets.

TABLE 1. INERT ORDNANCE TARGETS

Standard Type	Nonstandard (NS)
20-mm Projectile M55	20-mm Projectile M55
	20-mm Projectile M97
40-mm Grenades M385	40-mm Grenades M385
40-mm Projectile MKII Bodies	40-mm Projectile M813
BDU-28 Submunition	
BLU-26 Submunition	
M42 Submunition	
57-mm Projectile APC M86	
60-mm Mortar M49A3	60-mm Mortar (JPG)
	60-mm Mortar M49
2.75-inch Rocket M230	2.75-inch Rocket M230
	2.75-inch Rocket XM229
MK 118 ROCKEYE	
81-mm Mortar M374	81-mm Mortar (JPG)
	81-mm Mortar M374
105-mm Heat Rounds M456	
105-mm Projectile M60	105-mm Projectile M60
155-mm Projectile M483A1	155-mm Projectile M483A
	500-lb Bomb

JPG = Jefferson Proving Ground
HEAT = high-explosive, antitank

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

POC: John Breznick
(434) 978-3187
JBreznick@naevageophysics.com

Address: NAEVA Geophysics Inc.
P.O. Box 7325
Charlottesville, VA 22906

2.1.2 System Description (provided by demonstrator)

Dual EM61 MKII/towed:

This system will be employed to survey the Calibration Lanes, the Blind Test Grid, the Open Field Site, and the Desert Extreme Site. During the fall of 2003, NAEVA developed and field tested a new towed-array system for the Geonics EM61 MKII. Two 1- by 0.5-meter coils were encased in a durable polyplastic sled that rested directly on the ground. Coil heights can be adjusted using inflatable air bladders within the sled, but are typically maintained at the standard height of 40 cm above the ground, equivalent to mounting the coils on their standard wheels. The system is towed by an eight-wheeled Argo all-terrain vehicle. A 16-foot tongue attaches the coil assembly to the Argo and maintains sufficient separation so that the vehicle does not influence the geophysical data. A single Global Positioning System (GPS) sensor is mounted over the center of the two coils to provide real-time positional tracking capabilities. System electronics are securely mounted in the vehicle's rear compartment, and the data loggers are located in the driver's compartment to allow continuous monitoring of system function.

The system was designed with the goal of quickly collecting the highest quality geophysical data on a modular, reusable platform. The smooth-bottomed sled allows the system to negotiate rough terrain without the jarring and associated mechanical noise usually found in wheel-mounted systems. Lightweight and durable, the polyplastic shell is composed of several pieces that can be quickly replaced if field repairs are necessary. In addition, the coils are fully enclosed during operation, allowing the towed-array a degree of weatherproofing not usually found in geophysical equipment.

The EM61 is a time-domain electromagnetic instrument designed to detect, with high spatial resolution, shallow ferrous and nonferrous metallic objects. The applicability of the instrument for ordnance and explosives (OE) detection has been widely demonstrated at sites across the United States. Each instrument consists of two air-cored coils (1 by 0.5 m), batteries, processing electronics, and a digital data recorder. The larger of the two coils functions as the electromagnetic (EM) source and receiver and is positioned 40 cm below a second receiver coil. Secondary currents induced in both coils are measured in millivolts (mV).

Geonics has recently updated their standard EM61 system to the EM61 MKII. The primary difference in the MKII system is the use of multiple time gates. A time gate is the time after the electromagnetic pulse is generated that the receiver coil measures the response. The standard EM61 offers a single time gate in both the bottom and the top coils. While the top coil time gate is unchanged, the MKII records early, middle, and late channels from the bottom coil. The late time gate (third channel) corresponds to the standard EM61, and the earlier time gates offer enhanced capabilities for the detection of smaller metallic objects. Data from all three channels will be stored and processed during the demonstrations at APG.

Single EM61 MKII:

This system will be employed to survey the Calibration Lanes, the Blind Test Site, and the Mogul Challenge. In an effort to maintain the highest standards for quality data acquisition in an area suspected to have small munitions, the EM61 will be operated in a litter/strecher configuration, where the coils are supported by 12-foot-long fiberglass poles and transported by two operators (fig. 1). The data logger and backpack will be controlled by the operator at the back of the system. Coil height, consistent with the towed-array at 40 cm, will be maintained through the use of harnesses worn by both operators. NAEVA has found data quality in the tandem configuration to be superior to wheeled operation in all but the smoothest terrain.

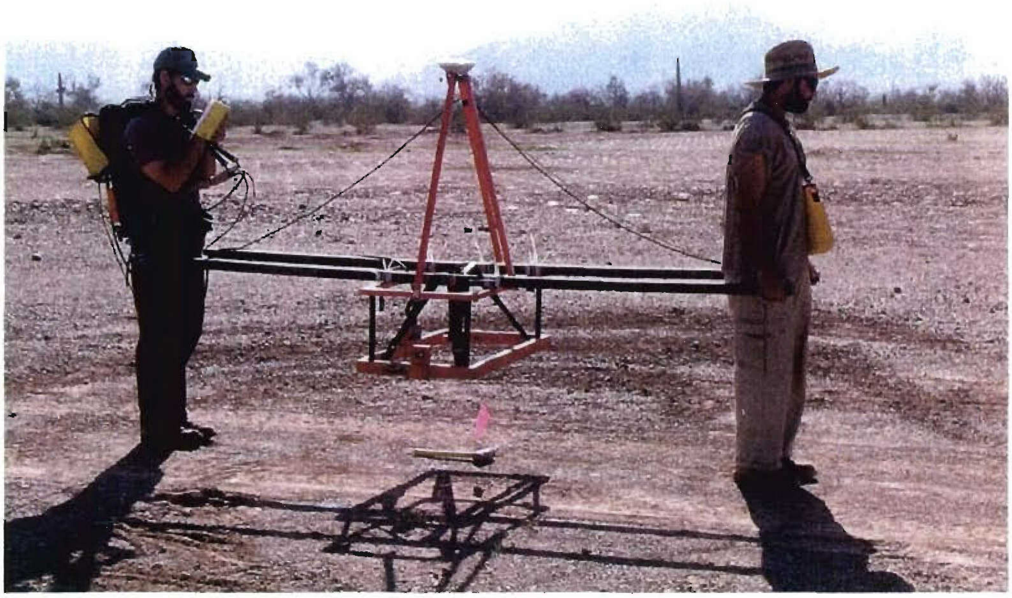

Figure 1. Demonstrator's system, EM61 MKII/man-portable.

2.1.3 Data Processing Description (provided by demonstrator)

All towed-array data will be collected with real-time GPS data positioning from an antenna mounted between the two coils. EM data will be collected at the rate of ten readings per second, which equates to more than one reading per foot. GPS locations will be logged at a rate of one reading per second. Real-time corrections from the GPS base receiver are broadcast to the roving GPS unit via a radio link. The GPS and electromagnetic data will be recorded in a single binary file on an Alegro field computer running Geonics' ML61MK2A software. This file is converted to a standard American Standard Code for Information Interchange (ASCII) file using Geonics' Multi61 Mark2 software. To maintain straight line profiling and to minimize the occurrence of gaps within the data, polyvinyl (PVC) pin flags will be used as ground control. The flags will be set in parallel lines across the area of investigation with alternating colors signifying the data collection paths. Pin flags will be spaced 8 feet apart, resulting in one pass with the array every 4 feet. Previous experience has shown that this spacing minimizes the occurrence of gaps between passes as well as provides overlapping coverage of the coil-to-coil gap inherent in the array. In addition, navigation and real-time field coverage will be aided by the use of StarPal software running on a Panasonic Toughbook computer linked to the GPS.

In areas of extremely rough terrain (Mogul Challenge), a single EM61 MKII will be hand-operated by field personnel. Data will be collected at the rate of 10 readings per second along lines spaced 2 feet apart. Raw binary data are collected on an Allegro portable field computer using EM61 MK2A Software. This file is converted to a standard ASCII file using Geonics' DAT61 MKII software.

Whether operating the towed-array or the hand-operated system, all geophysical mapping in open areas will make use of real-time GPS data positioning. In the case of the towed-array, the rover antenna will be mounted between the two coils and an offset will be applied during the post-processing to produce the actual coil positions. The rover antenna can be mounted directly over the single coil in hand-operated mode so that no offset is necessary.

If any areas are determined to have inadequate GPS satellite coverage, NAEVA will use tape measures and painted ropes to maintain accurate data positioning. Tape measures will be used with the existing control points to create a series of square grids to cover the area. Painted ropes will be placed every 25 feet, perpendicular to the direction of data collection. Evenly spaced painted marks on the ropes will allow the data collection team to maintain straight-line profiling over the area of investigation. Once all the data is collected, the control points will be used to transform the data from local coordinates to Geodetic Coordinates for scoring submittal. NAEVA has successfully used this method at numerous UXO sites where GPS coverage is not available.

The geophysical data will be temporarily stored in the instrument logger during data collection and then downloaded onto a laptop computer for on-site review and editing. Using Geosoft's Oasis Montaj software, a track plot of the instrument's GPS positions will be created to ensure that adequate data coverage has been achieved. For those areas without GPS coverage, Geonics' DAT61 MK2 software will be employed to correct the EM61 positioning using the fiducial marks entered in the data. Preliminary contour maps will then be created for field

review of each survey area. Once in-field processing and review are completed, the data will be electronically transferred to NAEVA's Virginia office for analysis and target selection.

Geosoft's Oasis Montaj UXO software package will be employed to post-process and contour the raw data, and to identify potential UXO targets. The program identifies peak amplitude responses of the frequency associated with, but not limited to, UXO items. Anomalies may generate multiple target designations depending on individual signature characteristics.

Geophysical data processing includes the following:

Instrument drift correction (leveling).

Lag correction.

Digital filtering and enhancement (if necessary).

Gridding of data.

Selection of all anomalies.

Selection of targets for intrusive characterization.

Preparation of geophysical and target maps.

Once NAEVA has completed the steps described above, the data will be forwarded to our subcontractor, AETC, for discrimination processing and final dig list development. AETC will evaluate only targets selected by NAEVA Geophysics. Their first step will be to invert the measured EM61 MKII data using a three-axis dipole model. AETC's EM61 fit algorithm determines the best set of induced dipole model parameters that account for the spatial variation of the EM61 signal as the sensor is moved over the object. The model parameters are target X,Y location and depth, three dipole response coefficients corresponding to the principle axes of the target, and the three angles that describe the orientation of the target. There is a set of three response coefficients for each of the EM61 MKII's four time gates. The magnitude of the response coefficients scales with the size of the target. An empirical relationship will be used to translate the sum of the target response coefficients into an equivalent UXO caliber. The relationship between the three response coefficients will tell us something about target shape. Cylindrical objects, like most UXO, have one large coefficient and two smaller, equal coefficients. Plate-like objects nominally have two large and one small coefficient.

Under controlled measurements, both the forward dipole model and fit algorithm have been found to be highly effective in describing EM61 measurements over buried ordnance. The accuracy of the fit algorithm has been found to be limited by poor quality data. In particular, closely spaced and accurately positioned measurements by the EM61 sensor are important for good fit results. Also, the model only describes the EM61 signal from compact objects and does not apply to extended objects such as utility lines.

2.1.4 Data Submission Format

Data were submitted for scoring in accordance with data submission protocols outlined in the Standardized UXO Technology Demonstration Site Handbook. These submitted data are not included in this report in order to protect ground truth information.

2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC) (provided by demonstrator)

Quality Control (QC):

To establish confidence in the data reliability, tests will be conducted in a systematic manner throughout the duration of the fieldwork. Various types of QC data are generated prior to, during, and after all data collection sessions.

Daily: A location identified as having no subsurface metal will be designated as a calibration point. Readings will be collected in a stationary position over the calibration point to ensure a stable and repeatable response was exhibited. During this time, a metallic item will be placed in a standard position with respect to the coils, and the instrument's response will be observed. The item will then be removed and static readings will continue. This test is performed daily to verify that the instrument is functioning properly, as indicated by a stable and repeatable response. The calibration point will also document the continued accurate performance of the GPS equipment.

A second location will be established over a buried item of known response, likely within one of the Calibration Lanes. At the start and end of each field day, two lines will be collected bidirectionally across the item along the same survey line. The data will then be reviewed for consistent response and positioning and to determine an appropriate lag correction.

During data collection: Upon completion of the original collection of a data set, approximately 3 percent of the line footage for each surveyed area will be recollected as a check of instrument repeatability and positioning. The repeat lines will be saved to separate files and used to create profiles that provide direct comparison with the original data. Each profile will be evaluated for repeatability in both instrument response and data positioning.

Overview of Quality Assurance (QA):

For purposes of this investigation, QA is defined as the procedures to be employed during the demonstration. All of the procedures are designed to provide excellent data quality while maximizing production during the field efforts.

All towed-array data will be collected with real-time GPS data positioning from an antenna mounted between the two coils. Electromagnetic data will be collected at the rate of 10 readings per second, which equates to more than one reading per foot. GPS locations will be logged at a rate of one reading per second. To maintain straight line profiling and to minimize the occurrence of gaps within the data, PVC pin flags will be used as ground control. The flags will be set in parallel lines across the area of investigation with alternating colors signifying the data

collection paths. Pin flags will be spaced 8 feet apart, resulting in one pass with the array every 4 feet. Previous experience has shown that this spacing minimizes the occurrence of gaps between passes as well as providing overlapping coverage of the coil-to-coil gap inherent in the array. While the GPS has a listed accuracy of 3 cm, the expected accuracy of resultant target selections is signified by a circle with a 1-foot radius around each target.

NAEVA's hand-operated system will use GPS for data positioning in areas such as the Mogul Challenge where satellite coverage is available. In such areas the data collection procedures will be identical to those described above with the exception that the line spacing will be reduced to 2 feet. In areas where GPS coverage is found to be inadequate, tape measures will be used in conjunction with the established control points to create a series of square survey cells to completely cover the area of investigation. Within each survey cell, data collection will be controlled using a series of marked survey ropes positioned at 25-foot intervals perpendicular to the survey line direction. Alternating color codes painted on the ropes at 2-foot intervals facilitate straight line profiling with the instrumentation during data collection. In addition, the ropes will serve as a point where the operator manually enters marks or fiducials into the data stream. The data will then be repositioned between the fiducials to account for the changes in velocity that occur as the instrument is carried across variable terrain conditions (i.e., slope, deadfall, vines, etc.). The inconsistent and difficult terrain expected at the site will dictate this relatively short fiducial separation (25 ft) to accommodate changes in velocity where greater care is necessary to navigate the instrument safely and effectively across the site.

2.1.6 Additional Records

The following record(s) by this vendor can be accessed via the Internet as MicroSoft Word documents at www.uxotestsites.org. The counterparts to this report are the Blind Grid, Scoring Record #666 and the Desert Extreme, Scoring Record #670.

2.2 YPG SITE INFORMATION

2.2.1 Location

YPG is located adjacent to the Colorado River in the Sonoran Desert. The UXO Standardized Test Site is located south of Pole Line Road and east of the Countermine Testing and Training Range. The Open Field range, Calibration Grid, Blind Grid, Mogul area, and Desert Extreme area comprise the 350 by 500-meter general test site area. The open field site is the largest of the test sites and measures approximately 200 by 350 meters. To the east of the open field range are the calibration and blind test grids that measure 30 by 40 meters and 40 by 40 meters, respectively. South of the Open Field is the 135- by 80-meter Mogul area consisting of a sequence of man-made depressions. The Desert Extreme area is located southeast of the open field site and has dimensions of 50 by 100 meters. The Desert Extreme area, covered with desert-type vegetation, is used to test the performance of different sensor platforms in a more severe desert conditions/environment.

2.2.2 Soil Type

Soil samples were collected at the YPG UXO Standardized Test Site by ERDC to characterize the shallow subsurface (< 3 m). Both surface grab samples and continuous soil borings were acquired. The soils were subjected to several laboratory analyses, including sieve/hydrometer, water content, magnetic susceptibility, dielectric permittivity, X-ray diffraction, and visual description.

There are two soil complexes present within the site, Riverbend-Carrizo and Cristobal-Gunsight. The Riverbend-Carrizo complex is comprised of mixed stream alluvium, whereas the Cristobal-Gunsight complex is derived from fan alluvium. The Cristobal-Gunsight complex covers the majority of the site. Most of the soil samples were classified as either a sandy loam or loamy sand, with most samples containing gravel-size particles. All samples had a measured water content less than 7 percent, except for two that contained 11-percent moisture. The majority of soil samples had water content between 1 to 2 percent. Samples containing more than 3 percent were generally deeper than 1 meter.

An X-ray diffraction analysis on four soil samples indicated a basic mineralogy of quartz, calcite, mica, feldspar, magnetite, and some clay. The presence of magnetite imparted a moderate magnetic susceptibility, with volume susceptibilities generally greater than 100 by 10-5 SI.

For more details concerning the soil properties at the YPG test site, go to www.uxotestsites.org on the web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at YPG is included in Table 2.

TABLE 2. ' TEST SITE AREAS

Area	Description
Calibration Grid	Contains the 15 standard ordnance items buried in six positions at various angles and depths to allow demonstrator equipment calibration.
Blind Grid	Contains 400 grid cells in a 0.16-hectare (0.39-acre) site. The center of each grid cell contains ordnance, clutter, or nothing.
Open Field	A 4-hectare (10-acre) site containing open areas, dips, ruts, and obstructions, including vegetation.
Mogul	A 2.64 acre area consisting of two areas (the rectangular or driving portion of the course and the triangular section with more difficult, non-drivable terrain). A series of craters (as deep as 0.91m) and trenches (as deep as 0.91m) encompass this section.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (7, 13, 14, and 16 December 2004)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total number of hours operated at each site are summarized in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	Number of Hours
Calibration Lanes	3.50
Mogul	10.66

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

A YPG weather station located approximately one mile west of the test site was used to record average temperature and precipitation on a half hour basis for each day of operation. The temperatures listed in Table 4 represent the average temperature during field operations from 0700 to 1700 hours while precipitation data represents a daily total amount of rainfall, if the data is not provided in the table below, the data was not taken by the YPG weather station. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 2004	Average Temperature, °F	Total Daily Precipitation, in.
7 December	NA	NA
13 December	57.83	0.00
14 December	60.42	0.00
16 December	NA	NA

3.3.2 Field Conditions

The field was dry and the weather warm throughout the NAEVA survey.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: Blind Grid, Calibration, Desert Extreme, Open Field areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are included in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization and daily equipment preparation and break down. A four-person crew took 55 minutes to perform the initial setup and mobilization. There was 1-hour and 50 minutes of daily equipment preparation and end of the day equipment break down lasted 45 minutes.

3.4.2 Calibration

NAEVA spent a total of 3 hours and 30 minutes in the calibration lanes, of which 1 hour and 45 minutes was spent collecting data. NAEVA also spent 40 minutes calibrating in the Mogul area.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, Demonstration Site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor costs (section 5) except for downtime due to Demonstration Site issues. Demonstration Site issues, while noted in the Daily Log, are considered non-chargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total Site Survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment data checks and maintenance activities accounted for 15 minutes of site usage time. These activities included changing out batteries and routine data checks to ensure the data was being properly recorded/collected. NAEVA spent an additional 1-hour and 20 minutes for breaks and lunches.

3.4.3.2 Equipment failure or repair. No time was needed to resolve equipment failures that occurred while surveying the Mogul.

3.4.3.3 Weather. No weather delays occurred during the survey.

3.4.4 Data Collection

NAEVA spent a total time of 10 hours and 40 minutes in the Mogul area, 6 hours and 30 minutes of which was spent collecting data.

3.4.5 Demobilization

The NAEVA survey crew went on to conduct a full demonstration of the site. Therefore, demobilization did not occur until 16 December 2004. On that day, it took the crew 2 hours and 10 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

NAEVA submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data was also provided within the required 30-day timeframe.

3.6 DEMONSTRATOR'S FIELD PERSONNEL

Leif Riddervold: Operations Manager
Alexander Kostera: General Field Support
Ashley Mowery: Towed Array System Operator
David Garey: Person Portable System Operator

3.7 DEMONSTRATOR'S FIELD SURVEYING METHOD

NAEVA surveyed the Moguls in a north to south direction.

3.8 SUMMARY OF DAILY LOGS

Daily logs capture all field activities during this demonstration and are located in Appendix D. Activities pertinent to this specific demonstration are indicated in highlighted text.

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL ORDNANCE CATEGORIES

Figure 2 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive. Figure 3 shows both probabilities plotted against their respective background alarm rate. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

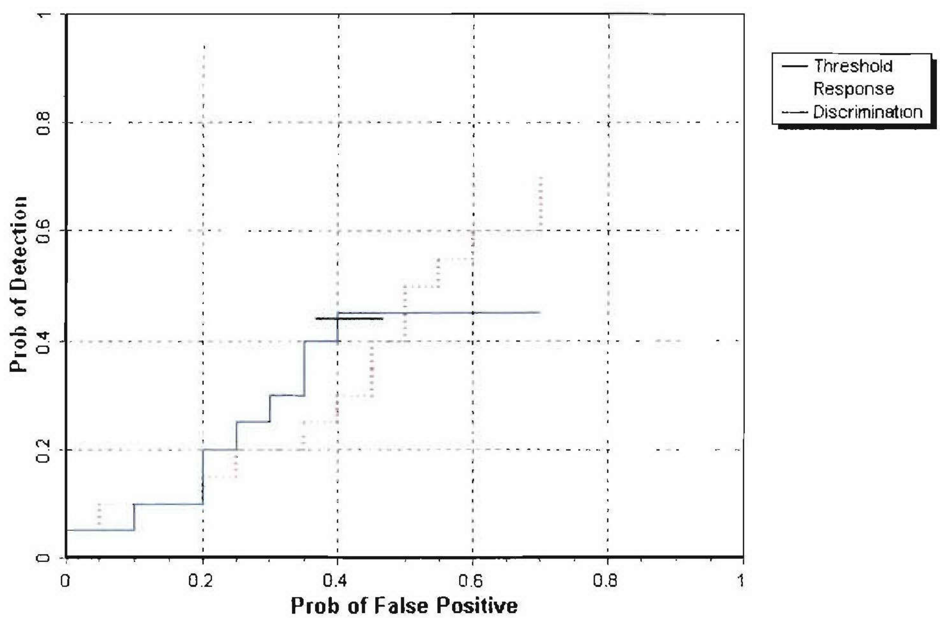

Figure 2. EM61 MKII/man-portable mogul probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

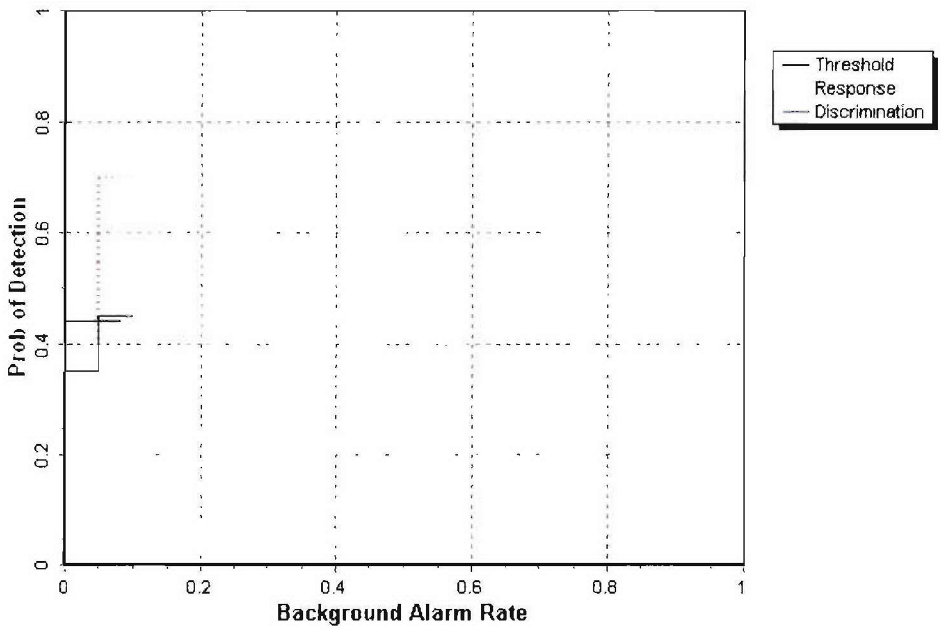

Figure 3. EM61 MKII/man-portable mogul probability of detection for response and discrimination stages versus their respective background alarm rate over all ordnance categories combined.

4.2 ROC CURVES USING ORDNANCE LARGER THAN 20 MM

Figure 4 shows the probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive when only targets larger than 20 mm are scored. Figure 5 shows both probabilities plotted against their respective probability of background alarm. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

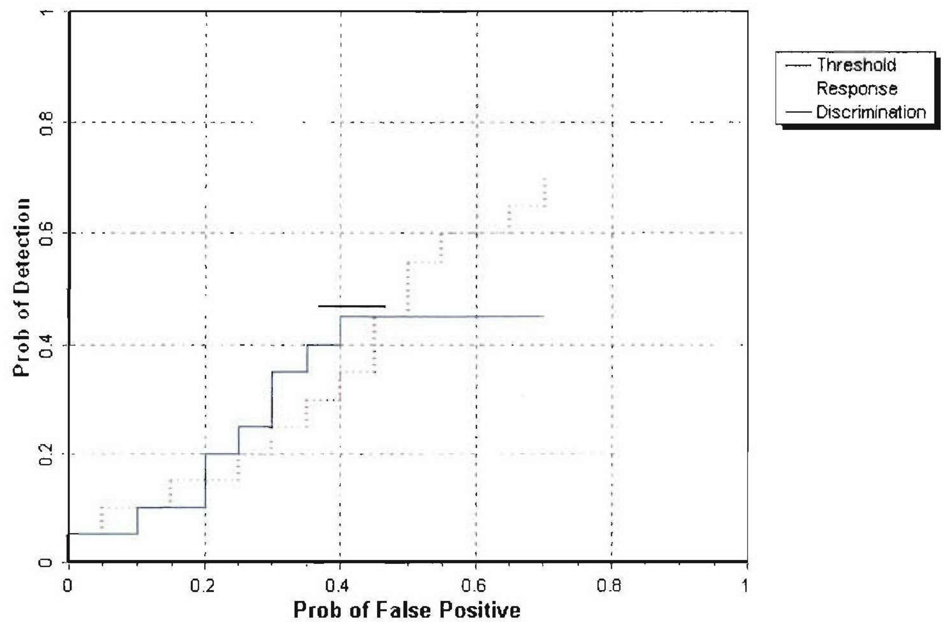

Figure 4. EM61 MKII/man-portable mogul probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

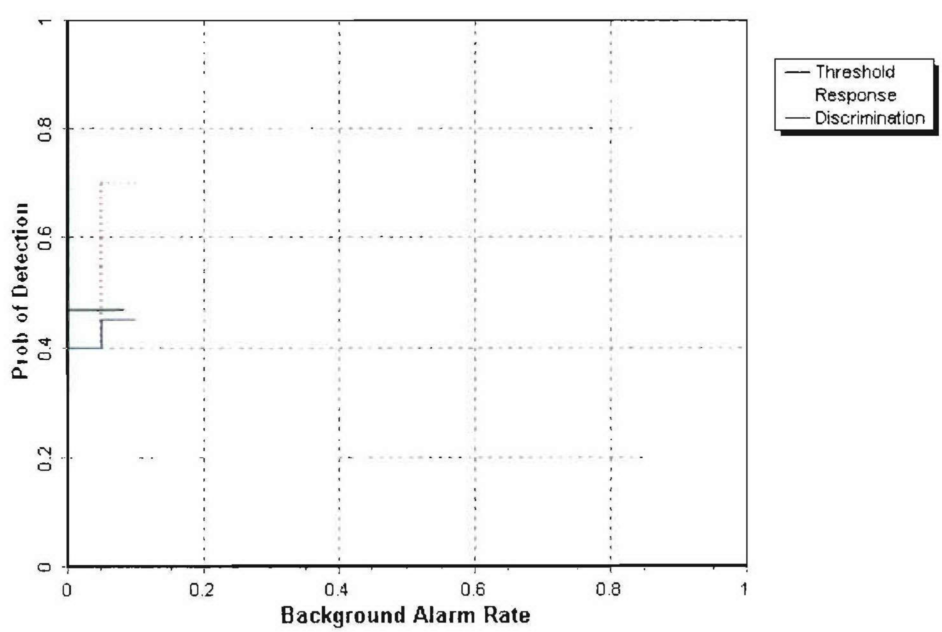

Figure 5. EM61 MKII/man-portable mogul probability of detection for response and discrimination stages versus their respective background alarm rate for all ordnance larger than 20 mm.

4.3 PERFORMANCE SUMMARIES

Results for the Mogul Area test, broken out by size, depth and nonstandard ordnance are presented in Table 5 (for cost results, see section 5). Results by size and depth include both standard and nonstandard ordnance. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range (see app A for size definitions). The results are relative to the number of ordnance items emplaced. Depth is measured from the geometric center of anomalies.

The RESPONSE STAGE results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the DISCRIMINATION STAGE are derived from the demonstrator's recommended threshold for optimizing UXO field cleanup by minimizing false digs and maximizing ordnance recovery. The lower 90 percent confidence limit on probability of detection and P_{fp} was calculated assuming that the number of detections and false positives are binomially distributed random variables. All results in Table 5 have been rounded to protect the ground truth. However, lower confidence limits were calculated using actual results.

TABLE 5. SUMMARY OF MOGUL RESULTS FOR EM61 MKII/MAN-PORTABLE

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.70	0.70	0.65	0.70	0.55	0.80	0.80	0.45	0.45
P_d Low 90% Conf	0.63	0.62	0.56	0.63	0.45	0.65	0.74	0.36	0.17
P_d Upper 90% Conf	0.74	0.77	0.76	0.79	0.68	0.93	0.87	0.59	0.72
P_{fp}	0.70	-	-	-	-	-	0.70	0.80	0.50
P_{fp} Low 90% Conf	0.66	-	-	-	-	-	0.63	0.71	0.05
P_{fp} Upper 90% Conf	0.75	-	-	-	-	-	0.73	0.89	0.95
BAR	0.10	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.45	0.45	0.40	0.55	0.30	0.40	0.50	0.30	0.30
P_d Low 90% Conf	0.38	0.39	0.30	0.44	0.20	0.25	0.44	0.20	0.08
P_d Upper 90% Conf	0.50	0.55	0.51	0.62	0.41	0.59	0.60	0.42	0.60
P_{fp}	0.40	-	-	-	-	-	0.40	0.60	0.00
P_{fp} Low 90% Conf	0.37	-	-	-	-	-	0.33	0.46	0.00
P_{fp} Upper 90% Conf	0.46	-	-	-	-	-	0.44	0.69	0.68
BAR	0.05	-	-	-	-	-	-	-	-

Response Stage Noise Level: 2.00
Recommended Discrimination Stage Threshold: 147.50

Note: The recommended discrimination stage threshold values are provided by the demonstrator.

4.4 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table 6.

TABLE 6. EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	0.64	0.41	0.60
With No Loss of P_d	1.00	0.03	0.01

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (table 7). Correct type examples include "20-mm projectile, 105-mm HEAT Projectile, and 2.75-inch Rocket". A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20mmP, 105H, and 2.75in, respectively.

TABLE 7. CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS UXO

Size	Percentage Correct
Small	2.9
Medium	18.2
Large	42.9
Overall	11.5

4.5 LOCATION ACCURACY

The mean location error and standard deviations appear in Table 8. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the Blind Grid, only depth errors are calculated, since (X, Y) positions are known to be the centers of each grid square.

TABLE 8. MEAN LOCATION ERROR AND STANDARD DEVIATION (M)

	Mean	Standard Deviation
Northing	0.02	0.16
Easting	-0.03	0.19
Depth	0.30	0.30

SECTION 5. ON-SITE LABOR COSTS

A standardized estimate for labor costs associated with this effort was calculated as follows: the first person at the test site was designated "supervisor", the second person was designated "data analyst", and the third and following personnel were considered "field support". Standardized hourly labor rates were charged by title: supervisor at $95.00/hour, data analyst at $57.00/hour, and field support at $28.50/hour.

Government representatives monitored on-site activity. All on-site activities were grouped into one of ten categories: initial setup/mobilization, daily setup/stop, calibration, collecting data, downtime due to break/lunch, downtime due to equipment failure, downtime due to equipment/data checks or maintenance, downtime due to weather, downtime due to demonstration site issue, or demobilization. See Appendix D for the daily activity log. See section 3.4 for a summary of field activities.

The standardized cost estimate associated with the labor needed to perform the field activities is presented in Table 9. Note that calibration time includes time spent in the Calibration Lanes as well as field calibrations. "Site survey time" includes daily setup/stop time, collecting data, breaks/lunch, downtime due to equipment/data checks or maintenance, downtime due to failure, and downtime due to weather.

TABLE 9. ON-SITE LABOR COSTS

	No. People	Hourly Wage	Hours	Cost
		Initial Setup		
Supervisor	1	$95.00	0.92	$87.40
Data Analyst	1	57.00	0.92	52.44
Field Support	2	28.50	0.92	52.44
SubTotal				**$192.28**
		Calibration		
Supervisor	1	$95.00	4.16	$395.20
Data Analyst	1	57.00	4.16	237.12
Field Support	2	28.50	4.16	237.12
SubTotal				**$869.44**
		Site Survey		
Supervisor	1	$95.00	10.66	$1012.70
Data Analyst	1	57.00	10.66	607.62
Field Support	2	28.50	10.66	607.62
SubTotal				**$2,227.94**

See notes at end of table.

TABLE 9 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
Demobilization				
Supervisor	1	$95.00	2.16	$205.20
Data Analyst	1	57.00	2.16	123.12
Field Support	2	28.50	2.16	123.12
Subtotal				**$451.44**
Total				**$3,741.10**

Notes: Calibration time includes time spent in the Calibration Lanes as well as calibration before each data run.

Site Survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to system maintenance, failure, and weather.

SECTION 6. COMPARISON OF RESULTS TO OPEN FIELD DEMONSTRATION

No comparison to date.

SECTION 7. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. If multiple declarations lie within R_{halo} of any item (clutter or ordnance), the declaration with the highest signal output within the R_{halo} will be utilized. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the length of the ordnance plus 1 meter.

Small Ordnance: Caliber of ordnance less than or equal to 40 mm (includes 20-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Ordnance: Caliber of ordnance greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75 in. Rocket, MK118 Rockeye, 81-mm mortar).

Large Ordnance: Caliber of ordnance greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, 155-mm projectile, 500-pound bomb).

Shallow: Items buried less than 0.3 meter below ground surface.

Medium: Items buried greater than or equal to 0.3 meter and less than 1 meter below ground surface.

Deep: Items buried greater than or equal to 1 meter below ground surface.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the Blind Grid test area.

Discrimination Stage Threshold: The demonstrator selected threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}) and those that do not correspond to any known item, termed background alarms.

The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the RESPONSE STAGE, the demonstrator provides the scoring committee with the location and signal strength of all anomalies that the demonstrator has deemed sufficient to warrant further investigation and/or processing as potential emplaced ordnance items. This list is generated with minimal processing (e.g., this list will include all signals above the system noise threshold). As such, it represents the most inclusive list of anomalies.

The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such, and to reject clutter. For the same locations as in the RESPONSE STAGE anomaly list, the DISCRIMINATION STAGE list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide "optimum" system performance, (i.e., that retains all the detected ordnance and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open Field only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to nonordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

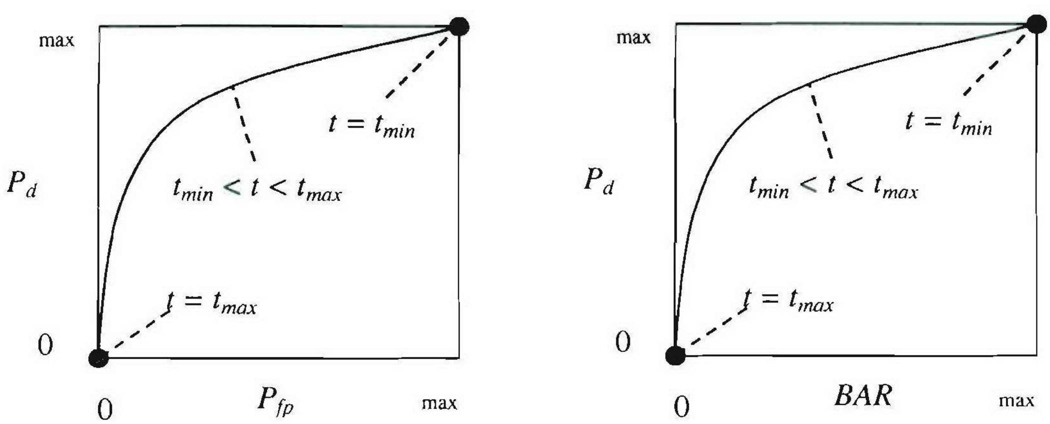

Figure A-1. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a pre-determined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the Blind Grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from nonordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$; Measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage tmin) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$; Measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage tmin). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind Grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open Field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})]$).

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 3).

A 2 x 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X's system is significantly degraded by the more challenging terrain feature introduced. The test statistic of the 2 x 2 contingency table is the

Chi-square distribution with one degree of freedom. Since an association between the more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A significance level of 0.05 is chosen which sets a critical decision limit of 2.71 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, the proportions are considered to be significantly different.

Standardized UXO Technology Demonstration Site examples, where blind grid results are compared to those from the open field and open field results are compared to those from one of the scenarios, follow. It should be noted that a significant result does not prove a cause and effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three progressively more difficult areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind Grid	Open Field	Moguls
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 ordnance out of 10 emplaced were correctly discriminated as such in open field-testing. Those four values are used to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 2.71, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{disc}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the smaller discrimination stage detection rate is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the ability of demonstrator X to correctly discriminate seems to have been degraded by the mogul terrain relative to results from the flat open field using the same system.

APPENDIX B. DAILY WEATHER LOGS

TABLE B-1. WEATHER LOG

Date	Time, HH:MM	Temperature (°C)	R/H, %	Precipitation, (in.)
12/06/2004	07:00	8.4	95	0.00
12/06/2004	08:00	8.1	95	0.00
12/06/2004	09:00	9.0	92	0.00
12/06/2004	10:00	11.2	77	0.00
12/06/2004	11:00	11.3	77	0.00
12/06/2004	12:00	13.2	69	0.00
12/06/2004	13:00	13.9	62	0.00
12/06/2004	14:00	10.7	57	0.00
12/06/2004	15:00	14.3	59	0.00
12/06/2004	16:00	14.9	55	0.00
12/06/2004	17:00	14.6	53	0.00
12/08/2004	07:00	6.0	95	0.00
12/08/2004	08:00	6.3	96	0.00
12/08/2004	09:00	8.4	89	0.00
12/08/2004	10:00	8.3	86	0.00
12/08/2004	11:00	6.2	81	0.00
12/08/2004	12:00	8.8	73	0.00
12/08/2004	13:00	13.2	65	0.00
12/08/2004	14:00	13.0	63	0.00
12/08/2004	15:00	13.5	63	0.00
12/08/2004	16:00	13.8	61	0.00
12/08/2004	17:00	12.7	63	0.00
12/09/2004	07:00	6.9	94	0.00
12/09/2004	08:00	6.7	95	0.00
12/09/2004	09:00	8.4	89	0.00
12/09/2004	10:00	10.5	82	0.00
12/09/2004	11:00	12.4	75	0.00
12/09/2004	12:00	13.8	68	0.00
12/09/2004	13:00	15.1	67	0.00
12/09/2004	14:00	-37.3	9	0.00
12/09/2004	15:00	-21.7	21	0.00
12/09/2004	16:00	-10.5	32	0.00
12/09/2004	17:00	5.0	47	0.00

TABLE B-1. (CONT'D)

Date	Time, HH:MM	Temperature (°C)	R/H, %	Precipitation, (in.)
12/10/2004	07:00	6.8	91	0.00
12/10/2004	08:00	4.8	92	0.00
12/10/2004	09:00	6.7	84	0.00
12/10/2004	10:00	11.1	72	0.00
12/10/2004	11:00	13.6	64	0.00
12/10/2004	12:00	15.3	60	0.00
12/10/2004	13:00	18.3	46	0.00
12/10/2004	14:00	20.0	42	0.00
12/10/2004	15:00	20.3	41	0.00
12/10/2004	16:00	20.1	39	0.00
12/10/2004	17:00	19.6	45	0.00
12/13/2004	07:00	5.7	90	0.00
12/13/2004	08:00	5.8	93	0.00
12/13/2004	09:00	8.6	89	0.00
12/13/2004	10:00	12.0	76	0.00
12/13/2004	11:00	13.9	71	0.00
12/13/2004	12:00	15.2	63	0.00
12/13/2004	13:00	17.1	51	0.00
12/13/2004	14:00	18.6	48	0.00
12/13/2004	15:00	18.2	46	0.00
12/13/2004	16:00	17.5	46	0.00
12/13/2004	17:00	17.2	45	0.00
12/14/2004	07:00	5.8	90	0.00
12/14/2004	08:00	5.8	90	0.00
12/14/2004	09:00	7.7	84	0.00
12/14/2004	10:00	12.8	66	0.00
12/14/2004	11:00	17.6	36	0.00
12/14/2004	12:00	19.3	27	0.00
12/14/2004	13:00	20.0	25	0.00
12/14/2004	14:00	20.4	24	0.00
12/14/2004	15:00	20.4	23	0.00
12/14/2004	16:00	20.8	23	0.00
12/14/2004	17:00	19.2	27	0.00

TABLE B-1. (CONT'D)

Date	Time, HH:MM	Temperature (°C)	R/H, %	Precipitation, (in.)
12/15/2004	07:00	4.8	89	0.00
12/15/2004	08:00	3.8	90	0.00
12/15/2004	09:00	6.3	80	0.00
12/15/2004	10:00	10.9	66	0.00
12/15/2004	11:00	13.8	54	0.00
12/15/2004	12:00	15.4	52	0.00
12/15/2004	13:00	17.3	43	0.00
12/15/2004	14:00	19.2	34	0.00
12/15/2004	15:00	15.8	31	0.00
12/15/2004	16:00	19.3	34	0.00
12/15/2004	17:00	19.0	35	0.00

APPENDIX C. SOIL MOISTURE

Date: 6 December 2004
Times: NA, 1300 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	NA	2.0
	6 to 12	NA	3.0
	12 to 24	NA	4.7
	24 to 36	NA	3.7
	36 to 48	NA	4.0
Mogul Area	0 to 6	NA	1.7
	6 to 12	NA	2.8
	12 to 24	NA	4.5
	24 to 36	NA	3.8
	36 to 48	NA	3.9
Desert Extreme Area	0 to 6	NA	1.5
	6 to 12	NA	2.1
	12 to 24	NA	3.7
	24 to 36	NA	3.7
	36 to 48	NA	3.9

Date: 7 December 2004
Times: 0730 hours, 1300 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.8	1.8
	6 to 12	2.9	2.9
	12 to 24	4.7	4.7
	24 to 36	3.7	3.7
	36 to 48	4.0	4.0
Mogul Area	0 to 6	1.7	1.7
	6 to 12	2.7	2.7
	12 to 24	4.5	4.5
	24 to 36	3.8	3.8
	36 to 48	3.9	3.9
Desert Extreme Area	0 to 6	1.5	1.5
	6 to 12	2.0	2.0
	12 to 24	3.7	3.7
	24 to 36	3.7	3.7
	36 to 48	3.9	3.9

Date: 6 December 2004
Times: 0730 hours, 1300 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.9	1.9
	6 to 12	2.8	2.8
	12 to 24	4.7	4.7
	24 to 36	3.7	3.7
	36 to 48	4.0	4.0
Mogul Area	0 to 6	1.7	1.7
	6 to 12	2.8	2.8
	12 to 24	4.5	4.5
	24 to 36	3.8	3.8
	36 to 48	3.9	3.9
Desert Extreme Area	0 to 6	1.5	1.5
	6 to 12	2.1	2.1
	12 to 24	3.7	3.7
	24 to 36	3.7	3.7
	36 to 48	3.9	3.9

Date: 9 December 2004
Times: 0730 hours, 1315 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.9	1.9
	6 to 12	2.8	2.8
	12 to 24	4.6	4.6
	24 to 36	3.7	3.7
	36 to 48	4.0	4.0
Mogul Area	0 to 6	1.7	1.7
	6 to 12	2.8	2.8
	12 to 24	4.4	4.4
	24 to 36	3.8	3.8
	36 to 48	3.9	3.9
Desert Extreme Area	0 to 6	1.5	1.5
	6 to 12	2.1	2.1
	12 to 24	3.7	3.7
	24 to 36	3.7	3.7
	36 to 48	3.9	3.9

Date: 10 December 2004
Times: 0710 hours, 1200 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.8	1.8
	6 to 12	2.8	2.8
	12 to 24	4.6	4.6
	24 to 36	3.7	3.7
	36 to 48	4.0	4.1
Mogul Area	0 to 6	1.7	1.7
	6 to 12	2.8	2.8
	12 to 24	4.5	4.5
	24 to 36	3.8	3.8
	36 to 48	3.9	3.9
Desert Extreme Area	0 to 6	1.5	1.5
	6 to 12	2.1	2.1
	12 to 24	3.7	3.7
	24 to 36	3.7	3.7
	36 to 48	3.9	3.9

Date: 13 December 2004
Times: 0715 hours, 1300 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.9	1.9
	6 to 12	2.7	2.7
	12 to 24	4.5	4.5
	24 to 36	3.7	3.7
	36 to 48	4.1	4.0
Mogul Area	0 to 6	1.8	1.8
	6 to 12	2.7	2.7
	12 to 24	4.5	4.5
	24 to 36	3.8	3.8
	36 to 48	3.9	3.9
Desert Extreme Area	0 to 6	1.6	1.6
	6 to 12	2.1	2.1
	12 to 24	3.7	2.7
	24 to 36	3.7	3.7
	36 to 48	3.9	3.9

Date: 14 December 2004
Times: NA, 1300 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.8	1.8
	6 to 12	2.7	2.7
	12 to 24	4.5	4.5
	24 to 36	3.7	3.7
	36 to 48	4.1	4.0
Mogul Area	0 to 6	1.7	1.7
	6 to 12	2.8	2.8
	12 to 24	4.5	4.5
	24 to 36	3.8	3.8
	36 to 48	3.9	3.9
Desert Extreme Area	0 to 6	1.5	1.5
	6 to 12	2.1	2.1
	12 to 24	3.7	3.7
	24 to 36	3.7	3.7
	36 to 48	3.9	3.9

Date: 6 December 2004
Times: NA, 1300 hours

Probe Location:	Layer, in.	AM Reading, %	PM Reading, %
Calibration Area	0 to 6	1.8	1.8
	6 to 12	2.7	2.7
	12 to 24	4.5	4.4
	24 to 36	3.7	3.7
	36 to 48	4.1	4.1
Mogul Area	0 to 6	1.7	1.7
	6 to 12	2.8	2.8
	12 to 24	4.5	4.5
	24 to 36	3.8	3.8
	36 to 48	3.9	3.9
Desert Extreme Area	0 to 6	1.5	1.5
	6 to 12	2.1	2.1
	12 to 24	3.7	3.7
	24 to 36	3.7	3.7
	36 to 48	3.9	3.9

APPENDIX D. DAILY ACTIVITY LOGS

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Track Method= Other Explain	Pattern	Field Conditions	
12/07/2004	4	CALIBRATION LANES	0715	0810	55	SETUP/DAILY START/ STOP/CALIBRATION	SETUP MOBILIZATION	NA	NA	NA	CLOUDY	COOL
12/07/2004	4	CALIBRATION LANES	0810	0815	5	SETUP/DAILY START/ STOP/CALIBRATION	CALIBRATED SYSTEM	NA	NA	NA	CLOUDY	COOL
12/07/2004	4	CALIBRATION LANES	0815	0915	60	COLLECT DATA	COLLECTED DATA BIDIRECTIONAL WEST TO EAST	GPS	NA	LINEAR	CLOUDY	COOL
12/07/2004	4	CALIBRATION LANES	0915	0940	25	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/ CHECK	CHECKING DATA	NA	NA	NA	CLOUDY	COOL
12/07/2004	4	CALIBRATION LANES	0940	1000	20	BREAK/LUNCH	BREAK	NA	NA	NA	CLOUDY	COOL
12/07/2004	4	BLIND TEST GRID	1000	1025	25	SETUP/DAILY START/ STOP/CALIBRATION	SETUP MOBILIZATION SET UP SYSTEM BTG	NA	NA	NA	CLOUDY	COOL
12/07/2004	4	BLIND TEST GRID	1025	1155	90	COLLECT DATA	COLLECTED DATA BIDIRECTIONAL WEST TO EAST	GPS	NA	LINEAR	CLOUDY	COOL
12/07/2004	4	BLIND TEST GRID	1155	1205	10	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/ CHECK	CHECKING DATA	NA	NA	NA	CLOUDY	COOL
12/07/2004	4	BLIND TEST GRID	1205	1225	20	BREAK/LUNCH	LUNCH	NA	NA	NA	CLOUDY	COOL
12/07/2004	4	MOGUL	1225	1305	40	SETUP/DAILY START/ STOP/CALIBRATION	SETUP/ MOBILIZATION SET UP TEST AREA MOGUL	NA	NA	NA	CLOUDY	COOL
12/07/2004	4	MOGUL	1305	1415	70	COLLECT DATA	COLLECTED DATA BIDIRECTIONAL NORTH TO SOUTH	GPS	NA	LINEAR	CLOUDY	COOL
12/07/2004	4	MOGUL	1415	1425	10	BREAK/LUNCH	BREAK	NA	NA	NA	CLOUDY	WARM
12/07/2004	4	MOGUL	1425	1525	60	COLLECT DATA	COLLECTED DATA BIDIRECTIONAL NORTH TO SOUTH	GPS	NA	LINEAR	CLOUDY	WARM

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Track Method= Other Explain	Pattern	Field Conditions	
12/07/2004	4	MOGUL	1525	1535	5	SETUP/DAILY START/ STOP/CALIBRATION	CALIBRATED SYSTEM STATIC CHECK	NA	NA	NA	CLOUDY	WARM
12/07/2004	4	MOGUL	1535	1600	25	SETUP/DAILY START/ STOP/CALIBRATION	END OF DAILY OPERATIONS/ EQUIPMENT BREAKDOWN	NA	NA	NA	CLOUDY	WARM
12/13/2004	2	MOGUL	1100	1105	5	SETUP/DAILY START/ STOP/CALIBRATION	SETUP MOBILIZATION SET UP TEST AERA MOGUL	NA	NA	NA	SUNNY	WARM
12/13/2004	2	MOGUL	1105	1115	10	SETUP/DAILY START/ STOP/CALIBRATION	CALIBRATED SYSTEM STATIC CHECK	NA	NA	NA	SUNNY	WARM
12/13/2004	2	MOGUL	1115	1225	70	COLLECT DATA	COLLECTED DATA BIDIRECTIONAL NORTH TO SOUTH	GPS	NA	LINEAR	SUNNY	WARM
12/13/2004	2	MOGUL	1225	1240	15	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/ CHECK	CHECK DATA	NA	NA	NA	SUNNY	WARM
12/13/2004	2	MOGUL	1240	1330	50	BREAK/LUNCH	LUNCH	NA	NA	NA	SUNNY	WARM
12/13/2004	2	MOGUL	1330	1525	115	COLLECT DATA	COLLECTED DATA BIDIRECTIONAL NORTH TO SOUTH	GPS	NA	LINEAR	SUNNY	WARM
12/13/2004	2	MOGUL	1525	1540	15	SETUP/DAILY START/ STOP/CALIBRATION	CALIBRATED SYSTEM STATIC CHECK	NA	NA	NA	SUNNY	WARM
12/13/2004	2	MOGUL	1540	1600	20	SETUP/DAILY START/ STOP/CALIBRATION	END OF DAILY OPERATIONS/ EQUIPMENT BREAKDOWN	NA	NA	NA	SUNNY	WARM
12/14/2004	2	MOGUL	0710	0815	65	SETUP/DAILY START/ STOP/CALIBRATION	SETUP MOBILIZATION SET UP TEST AREA MOGUL	NA	NA	NA	CLEAR	COOL
12/14/2004	2	MOGUL	0815	0825	10	SETUP/DAILY START/ STOP/CALIBRATION	CALIBRATED SYSTEM STATIC CHECK	NA	NA	NA	SUNNY	COOL

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

APPENDIX E. REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

4. Yuma Proving Ground Soil Survey Report, May 2003.

5. Practical Nonparametric Statistics, W.J. Conover, John Wiley & Sons, 1980, pages 144 through 151.

APPENDIX F. ABBREVIATIONS

AEC	=	U.S. Army Environmental Center
APG	=	Aberdeen Proving Ground
ASCII	=	American Standard Code for Information Interchange.
ATC	=	U.S. Army Aberdeen Test Center
EM	=	electromagnetic
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
EQT	=	Army Environmental Quality Technology Program
GPS	=	Global Positioning System
HEAT	=	high-explosive, antitank
JPG	=	Jefferson Proving Ground
OE	=	ordnance and explosives
POC	=	point of contact
PVC	=	polyvinyl chloride
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver-operating characteristic
SERDP	=	Strategic Environmental Research and Development Program
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground

APPENDIX G. DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-021

Addressee	No. of Copies
Commander U.S. Army Environmental Center ATTN: SFIM-AEC-ATT (Mr. George Robitaille) Aberdeen Proving Ground, MD 21010-5401	2
NAEVA Geophysics Inc. ATTN: (Mr. John Breznick) P.O. Box 7325 Charlottesville, VA 22906	1
SERDP/ESTCP ATTN: (Ms. Anne Andrews) 901 North Stuart Street, Suite 303 Arlington, VA 22203	1
Commander U.S. Army Aberdeen Test Center ATTN: CSTE-DTC-AT-SL-E (Mr. Larry Overbay)	1
(Library)	1
CSTE-DTC-AT-CS-R	1
Aberdeen Proving Ground, MD 21005-5059	
Defense Technical Information Center 8725 John J. Kingman Road, STE 0944 Fort Belvoir, VA 22060-6218	2

Secondary distribution is controlled by Commander, U.S. Army Environmental Center, ATTN: SFIM-AEC-ATT.

(Page G-2 Blank)